

METHOD FOR ENCODING AND DECODING MOVING PICTURE SIGNALS

50 BACKGROUND OF THE INVENTION

(1) Field of the Invention

The present invention relates to a method for encoding and decoding digital moving picture signals for use in TV phones, TV conferences and the like.

(2) Description of the Prior Art

In a general method for encoding digital moving picture signals, a frame of inputted moving picture is divided into plural blocks each composed of NxM pixels, and processes of motion detection, prediction, orthogonal transform, quantization, variable length coding, etc. are conducted on each block.

In a general method for decoding digital motion picture signals, blocks each composed of NxM pixels are regenerated in a reverse procedure, that is, processes of variable length decoding, reverse quantization, reverse orthogonal transform, motion compensation, etc.

The above general encoding method and decoding method for encoding and decoding digital moving picture signals enable removal of redundancy contained in moving picture signals, and efficient communication and storage of a moving picture with less information.

In the general encoding method and decoding method for encoding and decoding digital moving picture signals, the processes are conducted on each pixel block, as stated above. It is general that a set of pixel blocks forms a subframe and a set of subframes forms a frame, which are units processed in the general encoding and decoding method.

Hereinafter, encoding and decoding of each block, subframe and frame will be described by way of an example of a general encoding and decoding method for encoding and decoding digital moving picture signals with reference to ITU-T Recommendation H.261 (hereinafter, referred simply H.261) made on March, 1993.

H.261 defines an encoding method and a decoding method for encoding and decoding luminance signals and color difference signals, separately, of digital moving picture signals. However, description will be made of only the luminance signals, for the sake of convenience. Basically, the encoding method and decoding method for encoding and decoding the luminance signals are not different from those for the color difference signals.

As shown in FIG. 1, one frame 101 of digital moving picture signals is composed of 352x288 pixels according to H.261. The frame 101 is divided into twelve subframes 102 called GOBs (Group of Blocks) each composed of 176x48 pixels (hereinafter, the subframe in the description of the prior art will be referred a GOB). Further, the GOB 102 (subframe) is divided into thirty three blocks 103 called macro blocks each composed of 16x16 pixels.

The encoding method according to H.261 defines that encoded information for one frame is corresponded to a spatial hierarchical structure such as the frame 101, GOBs 102 and macro blocks 103 described above, as shown in FIG. 2.

In FIG. 2, a part enclosed in a rectangle shows encoded information, and the number of coding bits is shown under each of the rectangles. In FIG. 2, arrows show linkages of the encoded information. A series of encoded moving picture signal sequences as this is called a bit stream 104.

In the bit stream 104 according to H.261 shown in FIG. 2, a part including all encoded information for one macro

block 103 is called a macro block layer 103S, a part including all encoded information for one GOB 102 is called a GOB layer 102S, and a part including all encoded information for one frame 101 is called a frame layer 101S.

Meanings of the encoded information in each of the layers shown in FIG. 2 are given below:
Frame Layer 101S

PSC (20 bits): a frame identifier 105; a unique code by which an encoding method can be always identified, expressed as "0000 0000 0000 0000 0001";

TR (5 bits): a frame number 106; indicating a time position in which this frame 101 should be displayed;

PTYPE (6 bits): frame type information 107; various information about the frame 101;

PEI (1 bit): extension data insertion information 108; a flag representing presence of following PSPARE 109;

PSPARE (8 bits): extension data; GOB layer 102S (subframe)

GBSC (16 bits): a GOB identifier 110; a unique code by which a decoding method can be always identified, expressed as "0000 0000 0000 0000";

GN (4 bits): a GOB number 111; indicating a spatial position of this GOB 102 within the frame 101;

GQUANT (5 bits): quantization characteristic information 112; indicating a quantization characteristic when a macro block 103 in the GOB 102 is encoded;

GEI (1 bit): extension data insertion information 113; a flag representing presence of following GSPARE 114;

GSPARE (8 bits): extension data 114.

Incidentally, the encoded information 115 of the macro block layer which is the lowest hierarchy in FIG. 2 is generated in the encoding method of motion detection, prediction, orthogonal transform, quantization, variable length coding, etc., as described before, whose coding bit number is not fixed. The number of coding bits of the macro block layer 103S, in general, increases if a spatial level of pixels included in the macro block 103 changes largely or a time level of pixels included in the macro block 103 having the same spatial positions changes largely. Such macro block 103 is, hereinafter, referred a macro block 103 which is difficult to be encoded.

To the contrary, if a level of pixels included in the macro block 103 is steady in relation to space and time, the number of coding bits of the macro block layer 103S remarkably decreases, or sometimes becomes zero. Such macro block 103 is hereinafter referred a macro block 103 which is easy to be encoded.

In the decoding method according to H.261, the PSC 105 which is an identifier of the frame layer 101S is first found out from the bit stream 104. Incidentally, in a state where a decodable code has been successfully found out it is said that synchronization is established. When the PSC 105 is found out from the bit stream and synchronization of the frame layer 101S is established, it can be identified that the bit stream 104 until the next PSC 105 appears is encoded information for one frame. Further, a time position in which the frame 101 composed of 352x288 pixels obtained by decoding the bit stream 104 for that one frame can be obtained by examining the frame number 106 following the PSC 105.

After the establishment of the frame layer, a GBSC 110 that is an identifier of the GOB layer 102S is found out from the following bit stream 104 in the encoding method according to H.261. When synchronization of the GBSC layer is established, it can be identified that the bit stream 104 until the next GBSC 110 appears is encoded information for one GOB 102. Further, a spatial position of the GOB 102

composed of 176x48 pixels obtained by decoding the bit stream 104 for that one GOB 102 in a frame 101, in which the GOB 102 should be placed, can be obtained by examining a GN 111 which is a GOB number following the GBSC 110.

In the decoding method according to H.261, a bit stream 104 of a following macro block layer 103S is decoded after the establishment of the GOB layer 102s. The decoding method of the macro block layer 103S is a procedure to regenerate a macro block 103 composed of 16x16 pixels in processes of variable length decoding, reverse quantization, reverse orthogonal transform, motion compensation, etc., as described before. It should be here noted that the macro block layer 103S has no unique code by which a decoding method can be always identified dissimilarly to the PSC 105 or BGSC 110, and encoded information of each macro block is composed of undefined length bits of a variable length code.

As shown in FIG. 3, in the GOB (subframe) layer 102S, the encoded information from the first macro block 115, to the thirty third macro block 115₃₃ is expressed as a series of variable length codes without a unique code. If decoding of the macro block encoded information is initiated from a point indicated by A in FIG. 3, and successively conducted in the order of the first, the second.... the nth.... the thirty third macro blocks, it is possible to regenerate all the macro blocks 103 in the GOB layer 102S. However, if the decoding of the macro block encoded information is initiated from a point indicated by B or C in FIG. 3, it is impossible to identify a point from which encoded information 115 of one macro block starts, which leads to a failure of establishing synchronization. In which case, the decoding and regenerating all macro blocks 103 become unfeasible until the next GBSC 110 appears. In other words, the GBSC 110 also represents a starting point of decoding the macro block layer 103S.

Finally, in the decoding method according to H.261, the GOB 102 which is a set of regenerated macro blocks 103 is placed in a spatial position within a frame 101 directed by GN 111, and the frame 100 which is a set of the regenerated GOBs 102 is placed in a time position directed by TR 106.

As above, it is possible to decode one frame 101 of digital moving picture correctly in relation to space and time according to H.261.

However, the above general method for encoding and decoding digital moving picture signals has a drawback that if a part of a bit stream 104 is lacking or an error occurs therein, it might be impossible to accurately decode all subframes (GOBs) 102 in relation to time until synchronization of the next frame layer 101S is established.

The reason of the above is that codes which can be identified at all times in the bit stream 104 are only the PSC 105 which is a frame identifier and the GBSC 110 which is a subframe identifier in the general decoding method. If a part of the bit stream 104 lacks or an error occurs therein, it is impossible to recover synchronization of the decoding until the next GBSC 110 appears so that the decoding becomes unfeasible. Even if the next GBSC 110 appears, the bit stream 104 of that subframe layer 102S cannot be correctly decoded in relation to time. This will be understood from FIG. 4.

FIG. 4 shows an example where the fifth GOB 102₅ in the nth frame 101_n through the sixth GOB 102₆ in the (n+1)th frame 101_{n+1} cannot be decoded in relation to time due to lacking portions or errors of the bit stream 104 occurring in burst. In this example, not only the PSC 105 corresponding to the (n+1)th frame in relation to time but also the following TR

106 are missed or in error; it is therefore possible to correctly decode the GOB 102₇ in relation to space by establishing synchronization from the GBSC 110 corresponding to the seventh GOB 102₇ in the (n+1)th frame 101_{n+1} in relation to time and decoding the following GN 111, but impossible to specify whether this GOB 102₇ positions in the nth frame or in the (n+1)th frame in relation to time.

In terms of decoding of the eighth GOB 102₈ through the twelfth GOB 102₁₂ in the (n+1)th frame in relation to time, it is impossible to specify whether these GOBs 102 position in the nth frame or in the (n+1)th frame in relation to time.

In consequence, if a part of the bit stream 104 is missed or an error occurs therein, it becomes impossible to correctly decode all GOBs 102 in relation to time until synchronization of the next frame layer 101₅ is established.

Further, the general method for encoding and decoding digital moving picture signals has another drawback that if the GOB 102 including a picture in motion in relation to time cannot be decoded, a picture quality of the reproduced picture is largely degraded.

This problem will be described in more detail with reference to FIG. 5. FIG. 5 shows one frame including decoded signals of a moving picture, where a figure is moving in the center of the frame. In FIG. 5, a part moving in relation to time is indicated by slanting lines, and the remaining part is a background which is still in relation to time. A scene like this is general in TV conferences, TV telephones or the like.

Referring to FIG. 5, considering that any one of the first GOB 102₁ through the fourth GOB 102₄ cannot be decoded. The first through fourth GOBs 102₁ through 102₄ include a picture still in relation to time. If the second GOB 102₂ cannot be decoded, for example, a skillful operation is conducted to substitute the second GOB 102₂ of the present frame 101 with the second GOB 102₂ of the preceding frame 101₁ in the decoding. With this operation, degradation of a picture quality in the second GOB 102₂ of the present frame 101 may be hardly detected.

However, it is a problem if decoding of the fifth through twelfth GOBs 102₅ through 102₁₂ shown in FIG. 5 cannot be decoded. The fifth through twelfth GOBs 102₅ through 102₁₂ include a picture moving in relation to time. This means, for example, that a picture in the ninth GOB 102₉ of the preceding frame 101₁ is largely different from the ninth GOB 102₉ of the present frame 101 in relation to time. If the decoding of the ninth GOB 102₉ is unfeasible, degradation of the picture quality of the ninth GOB 102₉ of the present frame 101 is obviously detected even if the skillful operation mentioned above is conducted in the decoding.

Accordingly, if decoding of GOB 102 including a picture moving in relation to time becomes unfeasible, a quality of a reproduced picture is largely degraded.

SUMMARY OF THE INVENTION

In the light of the above problems, an object of the present invention is to provide a method for encoding and decoding digital moving picture signals, which can appropriately decode subframes (GOBs) following a subframe in trouble in relation to time if a part of a bit stream is missing or an error occurs in the bit stream.

Another object of the present invention is to provide a method for encoding and decoding digital moving picture signals, which can suppress degradation of a reproduced picture to a small extent if decoding of a subframe (GOB) including a picture in motion in relation to time becomes unfeasible.

To accomplish the first object, the present invention is featured in that in the method for encoding and decoding

digital moving picture signals of this invention, time position information representing an order of displaying a subframe is attached to an identifier of the subframe by which the subframe is identified.

According to the method for encoding and decoding digital moving picture signals of this invention, time position information representing an order of displaying a subframe is attached to an identifier used to identify the subframe and the identifier of the subframe is encoded. It is therefore possible to decode subframes following a subframe in trouble appropriately in relation to time if a part of [bit stream] a bit stream is missing or an error occurs in the bit stream by using the time position information representing an order of displaying each of the subframes attached to an identifier used to identify the subframe

To accomplish the second object, the present invention is featured in that in the method for encoding and decoding digital moving picture signals of this invention, the number of blocks included in a subframe is varied according to a sum of quantities of generated information of the blocks included in the subframe so that each of all the subframes included in the frame has an equal sum of quantities of the generated information of the blocks included in the subframe.

According to the method for encoding and decoding digital moving picture signals of this invention, the number of blocks included in a subframe is varied according to a sum of quantities of generated information of the blocks included in the subframe so that each of all the subframes included in the frame has an equal sum of quantities of the generated information of the blocks included in the subframe. In consequence, a spatial size of each subframe is not fixed. A subframe including a block having a large number of coding bits is in a smaller size, whereas a subframe including a block having a small number of coding bits is in a larger size. It is therefore possible to suppress degradation of a reproduced picture even if decoding of a subframe becomes unfeasible since a subframe including a block which includes a motion in relation to time and is difficult to be encoded is in a smaller size in relation to space.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 shows units to be encoded in a general encoding method for encoding moving picture signals;

FIG. 2 shows a bit stream generated in the general encoding method for encoding moving picture signals;

FIG. 3 shows a GOB layer in the bit stream in FIG. 2 generated in the general encoding method for encoding moving picture signals;

FIG. 4 illustrates an effect of a lack or an error of a part of a bit stream occurring in the general encoding and decoding method for encoding and decoding moving picture signals;

FIG. 5 illustrates an effect of a lack or an error of a part of a bit stream occurring in the general encoding and decoding method for encoding and decoding moving picture signals;

FIG. 6 shows a bit stream generated in a method for encoding digital moving picture signals according to first and second embodiments of this invention;

FIG. 7 is a flowchart illustrating the method for decoding digital moving picture signals according to the first embodiment of this invention;

FIG. 8 illustrates the method for encoding digital moving picture signals according to the second embodiment of this invention; and

FIG. 9 shows a structure of subframes according to the second embodiment of this invention.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

Hereinafter, description will be made of embodiments according to the present invention referring to the drawings.

A method for encoding and decoding digital moving picture signals according to a first embodiment will be now described, which may correctly decode a subframe as a unit in relation to time even if a part of a bit stream is missing or an error occurs therein.

In the encoding method according to this embodiment, one frame of digital motion picture signals is composed of, for example, 352x288 pixels. The frame is divided into twelve subframes each composed of, for example, 176x48 pixels. Further, the subframe is divided into thirty three blocks 13 each composed of, for example, 16x16 pixels.

The encoding method according to this embodiment corresponds encoded information for one frame to a spatial hierarchical structure made up of a frame 11, subframes 12 and blocks 13 to generate a bit stream 14 as shown, for example, in FIG. 6.

Meanings of encoded information of each layer shown in FIG. 6 are given below: Frame layer 11S

FSC (20 bits): a frame identifier 15; a unique code by which a decoding method can be always identified, expressed as "0000 0000 0000 0001 0000";

Subframe Layer 12S

SFSC (16 bits): a subframe identifier 16; a unique code by which a decoding method can be always identified, expressed as "0000 0000 0000 0001";

SFNT (5 bits): a subframe time number 17; indicating a time position in which this subframe 12 should be displayed;

SFNS (4 bits): a subframe space number 18; indicating a spatial position in which the subframe 12 should be displayed;

SFQUANT (5 bits): quantization characteristic information 19; representing a quantization characteristic when a block 13 in the subframe 12 is encoded.

Incidentally, encoded information 20 in the block layer 13S which is the lowest hierarchy in FIG. 6 is generated in an encoding method of motion detection, prediction, orthogonal transform, quantization, variable length coding, etc., whose coding bit number are not fixed.

Now referring to FIG. 7, a decoding method according to this embodiment will be now described. First, an FSC 15 which is an identifier of a frame layer 11S is found out from a bit stream 14 to establish synchronization of the frame layer 11S.

After the establishment of synchronization of the frame layer 11S, an SFSC 16 which is an identifier of a subframe layer 12S is found out from the following bit stream 14 to establish synchronization of the subframe layer 12S. Then a subframe time number SFNT 17 and a subframe space number SFNS 18 following the SFSC 16 are examined. Next, a bit stream 14 of a block layer 13S is decoded. A method for decoding this block layer 13S is a procedure to regenerate the block in processes of, for example, variable length decoding, reverse quantization, reverse orthogonal transform, motion compensation, etc. Finally, the subframe 12 which is a set of the regenerated blocks 13 is placed in time and space positions instructed by the SFNT 17 and the SFNS 18. If synchronization of the decoding is lost due to a lack of a part of the bit stream 14 or an error therein, a seek for the SFSC 16 which is an identifier of the subframe layer

12S is started. A lack [for an error of] of a portion of, or an error in, the bit stream 14 can be detected from, for example, that a decoded value exceeds a range specified in advance or an unexpected code word appears when the bit stream 14 is decoded. When the SFSC 16 is detected and synchronization of the subframe layer 12S is established, the SFNT 17 and the SFNS 18 are examined as stated above, the block layer 13S is decoded and regenerated, and the subframe 12 which is a set of the regenerated blocks 13 is placed in time and space positions instructed by the SFNT 17 and the SFNS 18.

According to the first embodiment of this invention, if a part of the bit stream 14 [lacks] is lacking, or an error occurs in the bit stream 14, synchronization of the decoding is lost and the decoding becomes unfeasible, but correct decoding becomes possible immediately after a subframe 12 in trouble.

As having been described the above first embodiment by way of an example, it is alternatively possible that the frame 11, the subframe 12 and the block are in different sizes and shapes. A bit length of each encoded information may be different from that of the above encoded information, or the frame layer 19 may be omitted, in addition.

According to a second embodiment of this invention, description will be now made of a method for encoding digital moving picture signals which can suppress degradation of a reproduced picture to a small extent if a subframe including a picture moving in relation to time cannot be decoded. Incidentally, it is possible here to employ a decoding method similar to that of the first embodiment.

In the encoding method of this embodiment, one frame 11 of digital moving picture signals is composed of, for example, 352x288 pixels. The frame 11 is divided into blocks each composed of 16x16 pixels. In other words, one frame 11 is composed of 22 blocksx18 block lines 21. The block line 21 corresponds to the subframe 12 mentioned above.

In the encoding method of this embodiment, each block 13' is encoded from the uppermost block line 21, as shown in FIG. 8, to generate encoded information. The encoded information of each block 13' is generated in an encoding method of, for example, motion detection, prediction, orthogonal transform, quantization, variable length coding, etc., the number of coding bits of which is not fixed. More specifically, the number of coding bits of a block 13' which is difficult to be encoded is large, whereas the number of coding bits of a block 13' which is easy to be encoded is small. In the encoding method of this embodiment, a set of blocks 13 or 13' composes a subframe 12 (or a block line 21) which is a unit of encoding, but the number of blocks 13 or 13' included in one subframe 12 or 12' is not fixed.

A manner of generating a bit stream 14 in the encoding method of this embodiment and a structure of a subframe layer 12S will be now described with reference to FIG. 6. When one frame is encoded, an identifier of a frame layer is encoded, and an FSC 15 is placed in a bit stream 14. Next, the identifier of the subframe layer 12S, a time number and a space number of that subframe, and a quantization characteristic of that subframe are encoded together, and code words of an SFSC 16, an SFNT 17, an SFNS 18 and an SFQUANT 19 are placed in the bit stream 14. At the same time, block coding bit number integrated value B-add is set to zero. Following that, a block 13 is encoded and encoded information of the block 13 composed of variable codes is placed in the bit stream 14. Concurrently, the coding bit number B of this block 13 is added to B-add. Namely, an equation, $B\text{-add}=B\text{-add}+B$, is computed. Similarly, blocks 13 are encoded successively, encoded information 20 of each block 13 is placed in the bit stream 14, and a calculation

of $B\text{-add}=B\text{-add}+B$ is repeated each time. If the B-add exceeds a subframe interval SFD when encoding of a certain block 12 is completed, an identifier of the subframe, a time number and a space number of that subframe and a quantization characteristic of that subframe are encoded, and code words of an SFSC 16, an SFNT 17, an SFNS 18 and an SFQUANT 19 are placed in the bit stream 14. At the same time, a block coding bit integrated value B-add is set to zero. In other words, a new subframe layer 12S is started to be formed from that point.

The subframe interval SFD is set to, for example, 540 bits. Therefore, if one frame is encoded with, for example, 6400 bits in the encoding method of this embodiment, 12 subframes 12 exist in one frame since $6400/540=11.85$.

In the encoding method according to this embodiment, the number of blocks 13 included in a subframe 12 is varied according to a quantity of generated information of the blocks included in one subframe, thereby varying a spatial size of the subframe 12, as stated above. More specifically, a subframe 12 including a block which is difficult to be encoded becomes small, whereas a subframe 12 including a block 13 which is easy to be encoded becomes large. FIG. 9 shows an example of a structure of subframes formed in the encoding method of this embodiment.

According to [the] the second embodiment of this invention, a subframe 12 including a block 13 which contains a motion in relation to time, and is thus difficult to be coded is made smaller in relation to space. If such subframe 12 cannot be decoded, it is possible to suppress degradation of a quality of a reproduced picture to a small extent. In a region within a frame in which no motion in relation to time exists and degradation of the picture quality is hardly detected even if the decoding is unfeasible, a size of one subframe is large in relation to space, which allows a less volume of side information such as the subframe identifier SFSC 16, subframe number SFNT 16 and subframe number SFNT 17. This can prevent an encoding efficiency from being lowered.

As having been described the second embodiment by way of an example, it is alternatively possible that the frame 11, subframe 12 and the block 13 are in different sizes and shapes. It is also possible to employ values of a quantity of codes of one frame and a subframe interval SFD different from those employed in the above example.

As obvious from the above embodiments, this invention enables correct decoding of each subframe 12 as a unit in relation to time even if a part of the bit stream 14 is missing or an error occurs therein.

Further, according to this invention, it is possible to suppress degradation of a quality of the reproduced picture to a small extent if a subframe 13 including a block which is in motion in relation to time cannot be decoded.

Still further, in a region within a frame in which no motion in relation to time exists and degradation of a quality of the reproduced picture quality is hardly detected even if the decoding is unfeasible, side information of the region is allowed to be in a small volume so that it is possible to prevent an encoding efficiency from being lowered.

What is claimed is:

(Amended). A method for encoding digital motion picture signals of a frame, comprising the steps of:

- dividing said frame into plural blocks each including NxM pixels;
- forming a subframe composed of a set of said blocks, said subframe being a unit to be encoded;
- setting an identifier to said subframe to identify said subframe; [and]
- specifying a frame to which said subframe belongs by adding [to said identifier] time position information to said identifier, the time position information representing an order of displaying said subframe;

6